SYSTEM AND METHOD FOR THE
MANAGEMENT OF DISTRIBUTED PERSONALIZED INFORMATION

BACKGROUND OF THE INVENTION

Field of the Invention

[0001]   The present invention relates to the field of information management. More specifically, the present invention, in an exemplary embodiment, relates to a system and method of operation for personalization of categorization of information and summaries of larger documents.

Description of the Related Art

[0002]   Knowledge base management tools are similar to data-mining and search tools. Data mining focuses on so-called legacy data that is more transactional and financial in nature. For example, a massive database of sales figures can reveal trends or allow a user to "drill down" to specific territories, products, and customers. Search tools may ferret out information known to be possibly present, but do not necessarily reveal anything inherent in the data. Knowledge base management tools go a step further by enabling a user to collect and organize information, search for what the user needs, and share the user's findings with others.

[0003]   In the art, most data provision services – i.e., search or data retrieval services – provide a set interface to their data. A user may only manipulate that interface in a manner dictated by the provider. For example, a user wishing to access services such as the LEXIS/NEXIS service which is provided by a division of Reed-Elsevier, Inc. or WESTLAW operated by the West Group may use software that resides at a user computer and that has access to data either resident on a CDROM, a hard drive local to the user computer's environment, or

via a data communications network such as the Internet. Other users may access those services directly through the Internet. Many information sources do not provide a user level taxonomy at all, relying, if at all, on an Internet browser or other software to provide some user level utility to organize information, e.g. once data are accessed, storage of the data is limited by the browser or operating system to folders and searching within folders limited as well.

[0004]   A problem that exists with these access methods is that the means and ability to categorize the data accessed is dictated by the provider. For example, WESTLAW uses its proprietary head notes taxonomy that WESTLAW alone maintains and creates. Educational services such as ERIC provide a rigid classification system.

[0005]   For many users, these taxonomies are either inadequate because they are not tailorable to the user's specific needs or ignored because they represent a view of categorization with which the user is neither familiar nor inclined to learn or use. Moreover, given the level of current art searching methods, use of the provider supplied and imposed taxonomy is not as attractive for searching as it may have once been.

[0006]   However, many users would benefit from an ability to categorize and organize data in a manner comfortable to that user, and perhaps to that user alone. Such capability will aid that user in accessing the data, extracting information relevant to that user from that data, and later retrieving that information rapidly and cost-effectively. A problem with such categorization is that such an ability may incur large support costs or be impracticably unwieldy.

[0007]   As is known in the art, so called knowledge bases may comprise an indexed, searchable set of queries or frequently asked questions (FAQs) coupled with a search engine. Some methods proposed in the prior art deal with mining generalized sequential patterns from large databases of raw data, taking into account user specified constraints such as taxonomies.

United States Patent 5,742,811 issued to Agrawal, et al. for "METHOD AND SYSTEM FOR MINING GENERALIZED SEQUENTIAL PATTERNS IN A LARGE DATABASE" is illustrative.

[0008]   The prior art also has much written on queries of databases. United States Patent 5,826,260 issued to Byrd, Jr., et al. for "INFORMATION RETRIEVAL SYSTEM AND METHOD FOR DISPLAYING AND ORDERING INFORMATION BASED ON QUERY ELEMENT CONTRIBUTION" is illustrative and teaches an information retrieval system wherein a query issued by a user is analyzed by a query engine into query elements. After the query has been evaluated against the document collections, a resulting hit list is presented to the user, e.g., as a table. The presented hit list displays not only an overall rank of a document but also a contribution of each query element to the rank of the document. The user can reorder the hit list by prioritizing the contribution of individual query elements to override the overall rank and by assigning additional weight(s) to those contributions. However, the prior art has not adequately addressed using queries as a method of capturing, condensing, and presenting raw data or its summarization according to a user defined, user configurable taxonomy.

[0009]   In addition, the prior art teaches methods of summarization of raw data. United States Patent 5,918,240 issued to Kupiec, et al. for "AUTOMATIC METHOD OF EXTRACTING SUMMARIZATION USING FEATURE PROBABILITIES" is illustrative. Kupiec teaches a method of automatically generating document extracts that makes use of feature value probabilities generated from a statistical analysis of manually generated summaries to extract the same set of sentences an expert might. The method is based upon an iterative approach.   However, Kupiec does not disclose and does not teach, suggest, or motivate towards

use of a user tailorable taxonomy when doing its summarization or association of the generated summarization with one or more elements in the taxonomy.

[0010]   Accordingly, there is a need for a user configurable interface comprising raw data, information summarized and derived from the raw data, and a user defined and maintained taxonomy to organize the information.

BRIEF DESCRIPTION OF THE DRAWINGS

[0011]   These and other features, aspects, and advantages of the present invention will become more fully apparent from the following description, appended claims, and accompanying drawings in which:

[0012]   **Fig. 1** is a schematic view of an exemplary embodiment of the system of the present inventions;

[0013]   **Fig. 2** is a schematic object-like view of organization of table elements of an exemplary embodiment of the system of the present inventions;

[0014]   **Fig. 3** is an exemplary taxonomy screen form;

[0015]   **Fig. 4** is an exemplary data entry screen showing entry by authority for a legal system;

[0016]   **Fig. 5** is an exemplary data entry screen showing entry by taxonomy element for a legal system;

[0017]   **Fig. 6** is an exemplary screen showing a form useful for searching of the summaries and authorities;

[0018]   **Fig. 7** is an exemplary screen showing a form useful for entry and searching of questions and FAQs; and

[0019] **Figs. 8a** and **8b** are flowcharts of an exemplary embodiment of the system of the present inventions.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENT

[0020] In general, throughout this description, if an item is described as implemented in software, it can equally well be implemented as hardware. Further, although numerous examples herein are described in terms of legal research, the present inventions are not limited by these exemplary embodiments but instead are limited only by claims.

[0021] Referring now to **Fig. 1**, the system and method of the present invention are scalable, as that term is understood by those of ordinary skill in the database arts. Database 22 may reside solely at server 20 or be distributed among the various computers 10,20,30,40. A first portion 200 (shown in **Fig. 2**) of database 22 may comprise captured raw data, referred to herein as "authority" data, where these raw data are typically captured into authority table 22a, and a second portion 300 (shown in **Fig. 2**) of database 22 may comprise user manipulatable data, referred to herein as "summary" data, typically in summary table 12a associated with computer 10. As will be appreciated by those of ordinary skill in the computer arts, database 22 may comprise a single database such as database 22 at server 20 or be distributed to include database 12 at computer 10.

[0022] As will be appreciated by those of ordinary skill in the computer arts, general data may be thought of as "raw data," unprocessed and/or not reduced into a condensed form. Raw data that are condensed and/or organized according to some schema may be thought of as "information," able to be understood and acted upon. By way of example and not limitation, a legal decision issued in written form by a court may contain hundreds of words, some of which describe the background of the issues before that court, some of which synopsize prior case or

statutory law which may apply to the facts of the case, and some of which define and discuss the holding of the court on the issues before that court. In this example, all the words of the written decision may be thought of as "raw data" but the two or three sentences within the decision that define the holding of the court on an issue may be what a user wishes to know and use, i.e. the "information" of the decision for that user.

[0023]   By way of further example, text documents and webpages may also be raw data as may blueprints, wiring diagrams, or the like useful in building projects or maintenance and troubleshooting operations.

[0024]   In many situations, the same raw data may be "mined" or synopsized by different users for different information. By way of further example and not limitation, a report of weather trends may be considered "raw data." A first user may view the raw weather data for temperature trends important to a building industry. A second user may view the raw weather data for rainfall trends important to a farming industry. Each user therefore uses the same raw data to mine different information.

[0025]   As used herein, "authority" is understood to mean any source of information and may comprise at least one of text data, graphics data, audio data, video data, and the like, or combinations thereof. Authority data may comprise documents such as documents, images, and the like. Authority data may further comprise references to the actual data, such as a URL or other file link. In an exemplary legal application, the documents may further comprise case law files, statutes, and court rules. Accordingly, by way of example and not limitation, authority table 22a may comprise at least one of legal data, medical data, educational data, manufacturing data, scientific data, and/or entertainment data. In a preferred embodiment, the actual raw data are captured for inclusion in authority table 22a such as in a memo field as will be familiar to those

of ordinary skill in so-called x-Base database software. However, a pointer or other reference may be captured instead or in combination, such as a file location on server 20 or an Internet URL reference.

[0026]   In the preferred embodiment, each entry in each authority table 22a, described herein as an "authority," will have a unique identifier such as a primary key value. Further, in a preferred embodiment, each authority may be identified by one or more keywords that may reside in an authority keywords table 22b associated with backend portion 200 (shown in **Fig. 2**). Authority keywords table 22b may be accessed using table software object 220b (shown in **Fig. 2**). Each element of authority keywords table 22b has a link back to authority table 22a from which that keyword was derived and may further comprise so-called pinpoint descriptors to identify the precise location in the authority associated with or resident in authority table 22a from which the keyword was derived.

[0027]   By way of example and not limitation, a text file is captured into authority table 22a and may be parsed according to any of a number of equivalent methods to tokenize or otherwise summarize the text file. Each token or summary identifier identified may be captured into an element in an associated authority keyword table 22b where the element in that table 22b for that token comprises a word offset from a given, known starting point, e.g. the beginning of the text document, where that token exists. By way of further example and not limitation, a user may arbitrarily associate one or more authority keywords with a file captured into or identified in an authority table 22a, such as a geographic location in a graphic file or a time sequence identifier for audio or audiovisual data. Additionally, in some currently envisioned alternative embodiments, authorities may be translated into a different language upon or after capture. Additionally still, the translation may occur on the fly using any one of numerous equivalent

methods or software as will be familiar to those of ordinary skill in the software arts and translations may be kept in authority table 22a and linked to each other to facilitate use.

[0028]   Accordingly, as discussed below, second portion 300 (shown in **Fig. 2**) of database 22 comprises user tailorable tables.   These tables comprise information and configuration data relevant to that user. In some instances the user tailorable tables may be shared between users, and in other instances the user tailorable tables may be distinct and not shared. By way of further example and not limitation, mechanics diagnosing engine problems for automobiles may wish to have access to the repair manuals for those automobiles as will those who repair air conditioner systems for those automobiles, although each of them will usually be interested in different subsets of the data contained in those repair manuals.

[0029]   Although database 22 and its first portion 200 and second portion 300 (both shown in **Fig. 2**) may be two or more databases, as that term is understood by those of ordinary skill in the database arts, in this document database 22, including its first portion 200 and its second portion 300, will be referred to as a single database. In an exemplary embodiment where the user and backend portions of the database co-reside, database 22 may be a single database. Further, in some currently envisioned embodiments, authority data may be raw data and not resident in a table. By way of example and not limitation, in an embodiment where the present invention is configured at least partially as an Internet browser with some portion of the authority data being accessible via the Internet such as via HTML or XML web pages using numerous means including Internet search engines.

[0030]   All of the software (not shown in **Fig. 1**) as well as database 22 may reside at a single computer, e.g. database 62 at computer 60 resident on a hard drive or optical storage

device. Further, personal digital assistant 50 may have database 22 such as a solid state memory device insertable into and useable by personal digital assistant 50.

[0031] In an alternative embodiment, a user or client portion of the software may reside at user computer 10, laptop 30, or personal digital assistant 50 with a backend portion remaining at server 20. In this scenario, user computer 10, laptop 30, or personal digital assistant 50 may communicate with server 20 via wired or wireless networks such as local area network 110 and/or via the Internet 100. In a similar manner, a portion of database 22 may be present remotely from server 20 such as database 12 or database 42. These other databases 12,42 may be separate from database 22 and may remain independent from or linked to database 22 by methods that will be familiar to those of ordinary skill in the database arts.

[0032] A backend software component (not shown in **Fig. 1**) may be present to handle access to database 22. Additionally, in its various embodiments, a user/client portion of the software (not shown in **Fig. 1**) may be present to provide an interface to database 22 that remains constant with access remaining transparent to the user. In a currently preferred embodiment, access to database 22 may be accomplished by using object oriented programming techniques in which data access is encapsulated in table software objects (generally referred to as "210" and "310" in **Fig. 2**) with a user front end programmed in C++ or HTML or the like. Each table software object, e.g. 12a, 22a, may therefore be abstracted at the user front end. These table software objects may handle accessing tables locally, via a local area network, via a wide area network, or a combination thereof, and use protocols such as data access objects (DAO), Universal Data Access (UDA), Remote Data Objects (RDO), ActiveX Data Objects (ADO), direct access, DCOM, ODBC, SQL, or the like, or a combination thereof, without having to change the user front end. Such abstraction, as will be familiar to those of ordinary skill in the

software programming arts, allows the present inventions to have an inherent scalability, as these terms will be familiar to those of ordinary skill in the software arts.

[0033]   As will be discussed herein below, the user interface allows a user to tailor various aspects of the user/client portion of the system to that user's desires and needs. In an embodiment, the inherent scalability of the present inventions will allow use of a browser or browser-like user front end exploiting HTML pages for menus and providing other user functions while using table software objects to access database 22 and backend functionality. Accordingly, in addition to the various functions described herein, back-end portion 200 (shown in **Fig. 2**) of the present inventions may be configured to allow the Internet to act as a portion of authority table 22a. Searches may therefore be accomplished using the search functions of the present inventions coupled with search facilities present via the Internet. Summaries may be generated from the Internet sources and stored locally according to the user's taxonomy.

[0034]   Further, the system may be used to provide selected users with access to one or more functions via the Internet. Using HTML, a user may sign in through an HTML user interface. In a currently preferred embodiment, the HTML user interface receives a temporary "cookie" from server 30 which is good only for the current session and is destroyed or otherwise deactivated when the current session ends or a predetermined inactivity timeout has occurred. Additionally, the sign-in process may require that the user first read and acknowledge a disclaimer similar to an end user license agreement. If the user rejects the disclaimer, a "cookie" will not be issued and, without the cookie, the user will be denied access to the system.

[0035]   Once access to the system has been granted, the user may be allowed access to predetermined system functionality such as to the frequently asked questions or in user question/FAQ table 12c. Should the user wish to access additional functionality, e.g. fee-based

functions, the user may then be presented with a request for payment such as with a form requesting payment via credit card (not shown in the figures). Once payment has been verified, the user may be issued an access code and password which will allow access to predetermined fee-based sections of the system. This access code and password may additionally be required for future access to the fee-based sections of the system. However, direct access to predetermined functionality may be disallowed until a user first accesses the system in the normal manner and acknowledges a disclaimer.

[0036]   In some currently envisioned embodiments, user anonymity is essential and may be assured by the use of one or more third party companies, such as those that handle credit card verification and subsequent funds transfer. In order to maintain anonymity, users to the additional functions may be assigned a discreet code such as an alphanumeric password code which shall be used as an identifier to track their queries and subsequent system responses. All communications between the system and these users may then occur through the user interface using this password code as the primary contact identifier. In certain embodiments, the password code may only be valid for the then current session, helping assure that no direct contact occurs between the a user and the system operators.

[0037]   Additionally, the present inventions may be accessible with or through additional software, by way of example and not limitation including operating systems, such as by a user or programmatically invocable service or a shortcut-key service, e.g. ALT-K; application systems, such as word processors, work flow processors, project planners, case mangers; and the like, or combinations thereof. Such embodiments may allow a user of another software application such as a word processor to utilize that other application while conducting searches and constructing documents or gathering other data from the system of the present inventions.

[0038]   As discussed above, in an alternative embodiment, the present inventions may be used in conjunction with a service that maintains authority table 22a. Server 30 is accessible to a user such as a laptop 30 or personal digital assistant 50 having appropriate access. As each such user researches or otherwise captures new authorities, server 30 may maintain either the authority raw data or a link to the authority raw data according to the teachings of the present inventions. If server 30 is accessible via the Internet 100, server 30 may be made available to a plurality of users on a general or selective basis, such as an open or free service, a for-fee service, a service limited to a community of users, and the like, or combinations thereof. Thus, research by one user, e.g. laptop 30, can inure to the benefit of all users, e.g. computer 10 and personal digital assistant 50, because authority table 22a grows with each user's additions and each authority so added may be made available such as via searches to all users. Although each user 30,50 may be allowed their own customized view of the authority data, e.g. their separate taxonomy tables 12a, all share the same source, i.e. authority table 22a.

[0039]   Referring now to **Fig. 2**, in the currently preferred embodiment, database 22 may comprise two major sections, 200 and 300. As will be familiar to those of ordinary skill in the database arts, a database in general may comprise tables and possibly other components such as views and reports. In this exemplary embodiment, backend portion 200 of database 22 comprises at least one authority table 22a, where authority table 22a may be accessible via table software object 210a, and user portion 300 comprises at least one summary table 12a, where summary table 12a may be accessible via table software object 310a. Both these tables may be free tables, associated with the same database such as database 22, or associated with different databases. However, in an exemplary embodiment where the user and backend functions co-reside in a computer, all tables may be associated with a single database such as database 22.

[0040] User portion 300 of database 22 comprises at least one user tailorable summary table 12b that may be accessible via table software object 310b and may comprise other user tailorable tables. In additional exemplary embodiments, one or more of these tables may be related to other entities, by way of example and not limitation including clients, client matters, resources, and the like, such as for use in billing, case management, work flow management, project planner, and the like, or combinations thereof.

[0041] User taxonomy table 12a may be part of user portion 300 and contains user defined categorizations as well as relationships between each member of taxonomy table 12a to others in taxonomy table 12a. In a currently preferred embodiment, user taxonomy table 12a is organized as a linked list where each taxonomy element, sometimes referred to herein as a "category," has a parent link value. However, individual elements of taxonomy 12a or entire groups of elements, e.g. branches of a hierarchically organized taxonomy table 12a, may be associated with or related to one or more other members or groups of taxonomy 12a as described more fully herein below.

[0042] User summary table 12b may be part of user portion 300 and contains an element for each item of information the user desires to link to at least one element in user taxonomy table 12a. In the currently preferred embodiment, each element in user summary table 12b, sometimes referred to herein as a "summary," must be linked to at least one element in user taxonomy 12a as well as to a single element in authority table 22a in backend portion 200. User summary table index 12f is associated with user summary table 12b in part to aid in searching functions. One currently envisioned method of relating user summary table 12b to authority table 22a is to use an intermediate table such as user knowledge table 12e which relates one

element in authority table 22a to one element in user summary table 12b, although one-to-many relationships may be defined in this or other manners as well.

[0043]   In an embodiment, user question/FAQ table 12c may exist in user portion 300. User question/FAQ table 12c may be used to contain organizational help queries or other user tailored condensations of data, e.g. serve as a container for a plurality of elements in user summary table 12b. Each element in user question/FAQ table 12c may be related to one or more entries in user knowledge table 12e. User knowledge table 12e, in turn, may be related to a plurality of elements in user summary table 12b. In alternatively contemplated embodiments, user question/FAQ table 12c and user knowledge table 12e may be combined into a single table. Use of separate tables allows a single element of user question/FAQ table 12c to have several elements in user knowledge table 12e related to that element in user question/FAQ table 12c, allowing a one-to-many relationship.

[0044]   In a further aspect of the present inventions, user question/FAQ table 12c may be reproduced as a report on a taxonomy element basis, e.g. with report table 12d. In this manner, summaries captured such as in user summary table 12b may be used in conjunction with user question/FAQ table 12c and taxonomy table 12a to create a user-based cylopedia or reference work.

[0045]   Because user portion 300 may be separate from backend portion 200, a single backend authority table 22a may be shared concurrently by numerous users, as will be familiar to those of ordinary skill in the database arts. Additionally, backend authority table 22a may reside at some distance remote from the user. In such an environment, user portion 300 software may communicate with backend portion 200 services, such as via wireless or wired networks such as local area network 110 (**Fig. 1**) or the Internet 100 (**Fig. 1**). Linkage between backend portion

200 and user portion 300 may be by a unique key identifier such as to a unique authority table key value. Additionally, messaging between backend portion 200 and the user portion 300 may comprise short messages, easing communications bandwidth requirements.

[0046]	In a currently envisioned embodiment, access to the system may be through numerous means, by way of example and not limitation including programs written for a specific database such as MICROSOFT VISUAL FOXPRO, general languages such as C++ or C# that may also include an SQL or ODBC or equivalent interface ability, an HTML (hypertext markup language) or the like graphical interface, JAVA, or the like, or combinations thereof. The user interface may thus be presented using an Internet browser or the like and may include access through the Internet. Additionally, one or more system functions may be executed from a persistent data storage device such as a hard disk, diskette, CDROM, e-prom, solid state device, or any other digital medium which can contain executable code and/or one or more portions of database 22.

[0047]	Referring now to **Fig. 3**, in a preferred embodiment, logic in forms (such as from 350 shown in **Fig. 3**) used for presentation of data to the user may be decoupled from data access logic by use of table software objects such as 210, 310 to allow adapting a resource being accessed so that the type of resource can change easily and independently.

[0048]	The user may acquire or otherwise access raw data for storage in authority table 22a (**Fig. 1**) which may reside at a backend server such as server 30 (**Fig. 1**). By way of example and not limitation, the user may copy these files from another source such as through the Internet 100 (**Fig. 1**) or directly enter the authority information into authority table 22a. In an exemplary embodiment, raw data may be input into authority table 22a by selecting an option, in **Fig. 3**

shown as button 355 with a label of FULL TEXT, which then allows the user to copy the data or link into authority table 22a.

[0049] Additionally, users may be allowed to switch between databases 22 as well. When a user switches, the context of the user's summaries and taxonomies may switch as well. However, a user may taxonomy table 12a between multiple authority tables 22a, as well as multiple tables in user portion 300. In situations where a user computer has or has access to multiple databases 22, if the user deletes a taxonomy level in such a multiple context the deleted item may remain present in taxonomy table 12a but would be disabled, e.g. grayed out, for the set of tables associated with the context in which the user deleted the item from taxonomy table 12a. If the item from taxonomy table 12a is associated with no further tables, that item may then be removed from the shared taxonomy table 12a.

[0050] Referring now to **Fig. 4**, once raw data are entered into or accessed from authority table 22a, a user may generate one or more summaries for each element of authority table 22a in numerous equivalent ways as will be familiar to those of ordinary skill in the programming arts. However, a user may invoke an entry screen such as shown at 400 and use that screen to capture or access an authority which has not yet been captured. If the authority is not yet present in authority table 22a, capturing the authority in this manner will, upon acceptance by the user such as with a button, capture that authority or its link into authority table 22a.

[0051] By way of example and not limitation, the user may retrieve the previously entered element from authority table 12a, view the raw data associated with that element, highlight or mark a section of the raw data to be captured, and signal a desire to capture the highlighted or marked section as a summary. This may be accomplished, by way of example and not limitation, by using a mouse to highlight the section and then using a right-button click of the

mouse to bring up a menu that has an entry to allow capture. In this manner, for example, a capture function could additionally capture the reference parameters such as word offset information for captured text data. The user may additionally associate the captured summary with one or more elements in taxonomy table 12a. Optionally, the user may create new elements in user taxonomy table 12a for use with the summary. The user may additionally edit the summary data captured and/or replace or augment all or a portion of the summary with whatever the user wishes, by way of example and not limitation including audio augmentation, video augmentation, and the like.

[0052]   Summary information may be shown such as at 420 and taxonomy elements associated with that summary for the authority may additionally be shown such as at 430. As shown in **Fig. 4**, an exemplary legal system form, additional information for the authority as well as the summary and taxonomy associations may be manipulated.

[0053]   Accordingly, the user can synopsize or summarize the raw data into information relevant to that user. In an additionally envisioned embodiment, an automatic summary option may be present to synopsize or suggest a synopsis of an authority or a portion of an authority. Further, a user may use form 400 to search for authorities and their summaries.

[0054]   Referring now to **Fig. 5**, the user may additionally search for and/or enter summary and/or authority data on a taxonomy element basis. As with **Fig. 4**, data so captured may be manipulated by the user.

[0055]   Referring now to **Fig. 6**, when a user desires to retrieve information, the user may search for information in numerous ways. By way of example and not limitation, the user may elect to see all summaries the user has created and associated with one or more selected elements in taxonomy table 12a (**Fig. 2**), to search the user created summaries in summary table 12b (**Fig.**

2) for all instances of one or more keywords, or a combination thereof. The user may additionally elect to search authority table 22a (**Fig. 2**) by keyword, including limiting searches to one or more parameters as will be familiar to those of ordinary skill in the database arts.

[0056]   In all search situations, keywords that are generated may be displayed such as at 620 or via button 652 for a user and modified or deleted, and additional keywords entered, by users with appropriate permission. In a further embodiment, the user may elect to designate how the keywords are to be viewed for a search to produce a result. As will be familiar to those of ordinary skill in the programming arts, the user query may designate certain words as required, such as by using a "+" symbol, and words to be omitted, such as by using a "-" symbol. Additionally, the user may use proximity search designators including "within sentence," "within paragraph," and/or "within 'n' words of 'phrase'" designators. In advanced searches, a user may further be able to designate additional search conditions for one or more fields present in authority table 12a, e.g. dates, authors, type of authority, and the like.

[0057]   In an embodiment, the user may also set a threshold value for searches such as at 622 whereby a query containing a plurality of keywords for the search requires an authority or summary to possess at least that threshold of keyword occurrences to be considered a positive or relevant result. By way of example and not limitation, a user query may contain four keywords and have a threshold of fifty percent. Any summary or authority, depending on the search requested, that has at least two of the four keywords is then presented to the user as a positive or relevant search result. Others, e.g. with only one keyword, are not presented.

[0058]   In an exemplary embodiment, keywords for authority table 22a, summary table 12b, user question/FAQ table 12c, or combinations thereof may be generated automatically or configurably. By way of example and not limitation, when new raw data are added to authority

table 22a, the raw data may be parsed, keywords generated, and keywords entered into authority keyword table 22b. Additionally, when a summary element or question/FAQ element is created or modified, its associated text may be parsed and keywords may be added to or replaced in the appropriate keywords table such as 12f. Noise words may be configurably excluded from the keywords tables, e.g. common words such as "this," "that," "with," "of," "the," and "a."

[0059]   Additionally, when new raw data are added to authority table 22a (**Fig. 2**) or when new elements are added into user question/FAQ table 12c (**Fig. 2**), the generated keywords may be used to suggest one or more taxonomy categorizations to the user. By way of example and not limitation, software executing at user computer 10,30,50 (**Fig. 1**) may receive notice that a new entry has been added to authority table 22a (**Fig. 2**). That software may gather the keywords associated with the new element of authority table 22a (**Fig. 2**) and then, using the user's taxonomy table 12a (**Fig. 2**), gather keywords associated by the user with that user taxonomy table 12a (**Fig. 2**) such as on an element by element basis. For each element in the taxonomy, the software can determine if those keywords are also present in the new authority's entry at or above a predetermined, configurable threshold. If so, the software can suggest that the newly entered authority is appropriate for the user with respect to that taxonomy category and additionally allow the user to view the authority and create new summaries.

[0060]   When a query uncovers no results meeting the search criteria, a user may invoke an Internet browser or other communications interface, either native to the system such as in a window or external to the system such as a second application, and access other sources of authority data. If data are found in the new source, the data may be captured into authority table 22a (**Fig. 2**) in accordance with the teachings of these inventions. Additionally, a user may optionally capture merely the reference to the data, e.g. a URL.

[0061]   In some embodiments, a natural language query section may be present, e.g. at 610. In other currently contemplated embodiments, question area 610 may be linked to user question/FAQ table 12c (**Fig. 2**). In such embodiments, users who may not have access to the system of the present invention may submit questions and users with access may review each question submitted and select from the summaries uncovered those which the querying user thinks relevant to the question asked. These relevant entries may then be associated with the user question, such as with button 651. By way of example and not limitation, a non-querying user may submit a question to be researched such as via e-mail and a querying user research the question and provide an answer back to the non-querying user.

[0062]   Referring now to **Fig. 7**, in an additionally contemplated embodiment, a further exemplary form is shown at 700 to allow further manipulation of questions and their answers. Such as those captured into user question/FAQ table 12c. The answer to the question may be edited by a user with appropriate access permission such as in answer box 710.

[0063]   By way of further example and not limitation, when viewing the search results that used summary table 12b as the source of the search, in an exemplary embodiment the user may select the summary created such as with the form in **Fig. 6** such as by double-clicking the mouse on the summary or selecting an appropriate command button or other equivalent method and be presented with the raw data from which the summary was derived. Such selection may invoke an editor to allow the user to redact the selected summaries into a more specific, targeted, and/or otherwise tailored response to the question posed.

[0064]   In an further currently envisioned alternate embodiment, Internet browser functionality may be accessible from a user perspective to present the interface such as discussed above in **Fig. 3** through **Fig. 7**. By way of example and not limitation, a user selectable option

may be present to invoke one or more functions from within a browser. In additionally contemplated embodiments, one or more functions of the user front-end may be integratable into an Internet browser as plug-in. By way of example and not limitation, using HTML, a user interface may provide a "front end" through which a user accesses the system, e.g. accesses its databases and/or communicates with the system operators. In certain embodiments, the user interface may be the only conduit between the user and the system.

[0065]   **In the operation of an exemplary embodiment**, referring now to **Fig. 8**, a user creates a taxonomy to the user's liking at step 800. The user taxonomy is stored in user taxonomy table 12a. The taxonomy is free-form, and as shown in **Fig. 3**, in the currently preferred embodiment may be organized and presented as a linked list presenting a tree structure as will be familiar to those who use Microsoft WINDOWS® software such as the file manager interface. In other currently envisioned embodiments, user taxonomy table 12a may be organized in any manner that allows multiple levels of associations between an element and other elements in that table 12a and may be presented to the user in numerous ways as will be familiar to those of ordinary skill in the computer arts, by way of example and not limitation including as folders and trees.

[0066]   Category elements in taxonomy table 12a are left up to a user with adequate permission to create, modify, or delete those members and can therefore reflect the categorizations and associations desired by a user individually. Although user taxonomy tables 12a may be shared among groups of users, each user can possess their own user taxonomy table 12a.

[0067]   When creating or manipulating the taxonomy, in certain embodiments a user may additionally designate branches of the taxonomy, e.g. categories, as being associated, such as by

being either equivalent or related. One element or "branch" can be equivalent to and/or related to multiple other elements or "branches." The user can then use a single taxonomy table 12a but dynamically switch to an alternate presentation of the taxonomy encompassed in that taxonomy table 12a, such as by designating a category or branch as the temporary root of an alternate presentation of the taxonomy. By way of example and not limitation, a user may have categories as shown in Table 1 below for money:

LEVEL1	LEVEL2	LEVEL3	LEVEL4
US			
	currency		
		paper	
		coin	
			Copper
			Silver
		credits	
CANADA			
	currency		
		paper	
		coin	
			Copper
			Silver
			Gold

TABLE 1

If the branch "us\currency" has been set to be equivalent to the branch "canada\curency," a dynamic alternate presentation may appear as shown in Table 2:

LEVEL1	LEVEL2	LEVEL3	LEVEL4
Currency			
	paper		
		US + CANADA	
	coin		
		Copper	
			US + CANADA
		Silver	
			US + CANADA
		Gold	
			CANADA
	credits		
		US	

TABLE 2

Alternatively, if the "us\currency\coin" branch is set to be equivalent to the "canada\curency\coin" branch, a further dynamic alternate presentation may appear as shown in Table 3:

LEVEL1	LEVEL2	LEVEL3
coin		
	Copper	
		US + CANADA
	Silver	
		US + CANADA
	Gold	
		CANADA

TABLE 3

[0068]   If the "us\currency\coin" branch is set as related to the "canada\curency\coin" branch, a further dynamic alternate presentation may appear as shown in Table 4:

LEVEL1	LEVEL2	LEVEL3	LEVEL4
coin			
	US		
		Copper	
			US
		Silver	
			US
	CANADA		
		Copper	
			CANADA
		Silver	
			CANADA
		Gold	
			CANADA

TABLE 4

[0069]   As a further illustrative example, the following taxonomy as shown in Table 5 may be created by a user:

LEVEL1	LEVEL2	LEVEL3	LEVEL4
engines			
	Manufacturer1		
		gas	
			4
			6
			8
		diesel	
			8
	Manufacturer2		
		gas	
			4
			6
			8
			10
		Diesel	
			8

TABLE 5

If the user designates "Manufacturer1\gas" as equivalent to "Manufacturer1\gas," a dynamic

alternate presentation may appear as shown in Table 6:

LEVEL1	LEVEL2	LEVEL3
Gas		
	4	
		Manufacturer1+ Manufacturer2
	6	
		Manufacturer1+ Manufacturer2
	8	
		Manufacturer1+ Manufacturer2
	10	
		Manufacturer2

TABLE 6

Similarly, if the user designates "Manufacturer1\gas" as related to "Manufacturer1\gas," a

dynamic alternate presentation may appear as shown in Table 7:

LEVEL1	LEVEL2	LEVEL3
Gas		
	4	
		Manufacturer1
		Manufacturer2
	6	
		Manufacturer1
		Manufacturer2
	8	
		Manufacturer1
		Manufacturer2
	10	
		Manufacturer2

TABLE 7

[0070]   In all of these examples above, the initial taxonomy table 12a will not be altered

by these alternate presentations. However, using these dynamic alternate presentations, the user

can enter data and/or summaries in one category, e.g. "engines\manufacturer1\gas\4" and be able

to search and retrieve information from that category when investigating a class of similar

categories, e.g. 4 cylinder gasoline engines in general.

[0071] An exemplary embodiment for a legal environment will now be described. Users may create their taxonomy in the manner identified above. There is no predefined taxonomy required. Thus, a lawyer concerned with intellectual property may create a taxonomy that describes patent, trademark, anti-trust, trade secret, and copyright categories. A legal researcher concerned with real estate law may create a taxonomy that describes real property, personal property, government regulations, and contracts where contracts has further subcategory layers for sales and leases and government regulations has subcategories encompassing environmental law.

[0072] Case law, statutory law, contracts, opinions, regulations, and the like may be captured 810, such as through the Internet, into authority table 22a. Many of these files are available through court or other governmental sites and are in the public domain and other such documents may be available without copyright restriction. The user may additionally manually capture the relevant authority identifier 812, e.g. the volume/reporter/page/court/date information, or such information may be automatically or programmatically derived.

[0073] Once the raw data are captured and the description of the source of raw data obtained, these data may then be associatively stored in their respective tables 824, e.g. authority table 22a.

[0074] Once captured, the user may then retrieve the raw data 826 and create one or more summaries 830, if so desired, as described herein above. Because the categories are embodied in a user modifiable taxonomy table 12a, users are free to organize their summarization as they wish, in a manner and with a layout helpful to that specific user. Further, as the user portion 300 may be separate from the backend portion 200, users may freely create different user portions 300 without impacting other users' use of the backend portion 200 authority table 22a. If a

taxonomy entry does not exist 832 that adequately categorizes the summary created by the user, according to the user's wishes, the user may create a new taxonomy entry 833. The user may then associate the summary with the desired taxonomy entry 834 and with the raw data from which the summary was derived 838.

[0075]   Control to access of all or portions of database 22 or any of its functions including access to and manipulation of backend portion 200 or user portion 300 may be achieved by numerous methods as will be familiar to those of ordinary skill in the database and programming arts, by way of example and not limitation including password logins, cookies, and the like, or combinations thereof.

[0076]   By way of further example and not limitation, the present inventions may be used in conjunction with other free or for-fee services, such as those available through the Internet 100, such as by being a front end to those other services. Acting as an intermediary, the present inventions allow users to contain and refine knowledge and references uncovered during prior searches and integrate new information into the user's captured knowledge and references. By way of example and not limitation, in this manner a user may continue to use services such as LEXIS/NEXIS™ offered by REID ELSEVIER, INC. or WESTLAW™ offered by THE THOMSON GROUP but not be required to utilize the taxonomies used by those services. Similarly, users may access search engines available through the Internet 100 such as YAHOO™ or GOOGLE™ which either have their own taxonomy or offer no taxonomy, and integrate information uncovered with those search into the user's captured knowledge and references.

[0077]   Referring now to **Fig. 8b**, once created, users may freely search 840 either their own user portion 300 or authority table 22a or both using simple, word-oriented queries or any other query available to the user through database 22, by way of example and not limitation

including latent semantic indexing. In a distributed environment, a user query could be communicated to server 30 executing a query handler, and results messaged back to the user query handler.

[0078] These searches may be retained such as by capturing relevant search results 842 and the retained results further edited 844 if desired.

[0079] It is expressly understood that although the exemplary embodiments above have been described in terms of multiple tables, whether free tables or database tables, numerous equivalent structures, table definitions, and database definitions exist which will be equivalent to the structure, table definitions, and database definitions above. It is also expressly understood that although the exemplary embodiments above have been described in terms of a legal search system, the present inventions are not limited to legal searches.

[0080] It will be understood that various changes in the details, materials, and arrangements of the parts which have been described and illustrated above in order to explain the nature of this invention may be made by those skilled in the art without departing from the principle and scope of the invention as recited in the following claims.